Journal of Exposure Analysis and Environmental Epidemiology (2005) 15, 111–120

www.nature.com/jea

Reliability of a semi-quantitative method for dermal exposure assessment (DREAM)

BERNA VAN WENDEL DE JOODE,[a,b] JOOP J VAN HEMMEN,[a] TIM MEIJSTER,[b] VICKY MAJOR,[c,d] LESLIE LONDON[c] AND HANS KROMHOUT[b]

[a]*Department of Chemical Exposure Assessment, TNO Chemistry, 3700 AJ Zeist, The Netherlands*
[b]*Environmental and Occupational Health Division, Institute for Risk Assessment Sciences, Utrecht University, 3508 TD Utrecht, The Netherlands*
[c]*School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa*
[d]*Department of Health Sciences, Peninsula Technikon, Cape Town, South Africa*

Valid and reliable semi-quantitative dermal exposure assessment methods for epidemiological research and for occupational hygiene practice, applicable for different chemical agents, are practically nonexistent. The aim of this study was to assess the reliability of a recently developed semi-quantitative dermal exposure assessment method (DREAM) by (i) studying inter-observer agreement, (ii) assessing the effect of individual observers on dermal exposure estimates for different tasks, and (iii) comparing inter-observer agreement for ranking of body parts according to their exposure level. Four studies were performed in which a total of 29 observers (mainly occupational hygienists) were asked to fill in DREAM while performing side-by-side observations for different tasks, comprising dermal exposures to liquids, solids, and vapors. Intra-class correlation coefficients ranged from 0.68 to 0.87 for total dermal exposure estimates, indicating good to excellent inter-observer agreement. The effects of individual observers on task estimates were estimated using a linear mixed effect model with logged DREAM estimates as explanatory variable; ‘‘task’’, ‘‘company/department’’, and the interaction of ‘‘task’’ and ‘‘company/department’’ as *fixed* effects; and ‘‘observer’’ as a *random* effect. Geometric mean (GM) dermal exposure estimates for different tasks were estimated by taking the exponent of the predicted betas for the tasks. By taking the exponent of the predicted observer's intercept ($\exp^{\omega i}$), a multiplier (M_O) was estimated for each observer. The effects of individual observers on task estimates were relatively small, as the maximum predicted mean observers' multiplier was only a factor 2, while predicted GMs of dermal exposure estimates for tasks ranged from 0 to 1226, and none of the predicted individual observers' multipliers differed significantly from 1 (t-test $\alpha = 0.05$). Inter-observer agreement for ranking of dermal exposure of nine body parts was moderate to good, as median values of Spearman correlation coefficients for pairs of observers ranged from 0.29 to 0.93. DREAM provides reproducible results for a broad range of tasks with dermal exposures to liquids, solids, as well as vapors. DREAM appears to offer a useful advance for estimations of dermal exposure both for epidemiological research and for occupational hygiene practice.
Journal of Exposure Analysis and Environmental Epidemiology (2005) **15,** 111–120. doi:10.1038/sj.jea.7500369
Published online 7 April 2004

Keywords: *semi-quantitative dermal exposure, reliability, reproducibility, inter-rater agreement.*

Introduction and background

Both in occupational hygiene and epidemiology, it can be worthwhile to semi-quantitatively assess exposure to hazardous substances, resulting in a relative ranking of task or jobs. Semi-quantitative exposure assessment can be used to optimize measurement strategies for quantitative exposure assessment or to set priorities for control measures.

In occupational epidemiology there has been an increasing trend to use experts, such as occupational hygienists and other professionals, to assess exposure to chemical agents semi-quantitatively (Teschke et al., 2002), as measurements are often considered to be too expensive to collect, or measurement results are unavailable when exposure has to be assessed retrospectively.

One of the disadvantages of exposure assessment by experts is that their underlying decision-making process usually is a black box of unstructured opinions about (routes of) probable exposure (Cherrie et al., 1996; Kromhout, 2002). By structuring the experts' exposure assessment, for example, by developing an algorithm on the basis of determinants of exposure, estimates are expected to have an increased reproducibility being influenced less by the estimator's subjective opinion (Cherrie et al., 1996).

Semi-quantitative exposure assessment has generally aimed at assessing inhalation exposures and hardly at assessing dermal exposures (Kromhout et al., 1987; Hertzman et al., 1988; Hawkins and Evans, 1989; Teschke et al., 1989; Post et al., 1991; de Cock et al., 1996; Siematycki et al., 1997;

1. Abbreviations: ANP, antineoplastic drugs; CNC, computer numeric control; DREAM, dermal exposure assessment method; ICC, intra-class correlation coefficient; MP, metal parts; O, observer.

2. Address all correspondence to: Dr. H. Kromhout, Environmental and Occupational Health Division, Institute for Risk Assessment Sciences, Utrecht University, PO Box 80176, 3508 TD Utrecht, The Netherlands. Fax: +31-30-253-9499.
E-mail: h.kromhout@iras.uu.nl
Published online 7 April 2004

Benke et al., 1997; Cherrie and Schneider, 1999; Vermeulen et al., 2002; Walker et al., 2003). Even for contaminants for which dermal exposure is known to contribute significantly to internal dose (e.g., pesticides, polycyclic aromatic hydrocarbons, and polychlorinated biphenyls (PCB)), exposure assessment is often directed to inhalation exposures, resulting in inaccurate and imprecise exposure estimates (Vermeulen et al., 2002).

Valid semi-quantitative dermal exposure assessment methods (DREAMs), applicable for a broad range of agents, are practically nonexistent. For example, dermal exposure estimates of the expert system for exposure assessment EASE are known to be inaccurate and imprecise (Cherrie et al., 2003), while other methods are only applicable for specific exposures and tasks like spray painting (Brouwer et al., 2001).

We therefore developed a DREAM, an observational semi-quantitative method to assess dermal exposures by systematically evaluating exposure determinants using pre-assigned default values (Van-Wendel-de-Joode et al., 2003). The method is supposed to be generic and to be used in occupational hygiene and epidemiology. The outcome is a numerical estimate, indicating the amount of dermal exposure workers encounter when performing a certain task or job.

DREAM supplies estimates for exposure levels on outside clothing layers (*potential dermal exposure*) as well as on skin (*actual dermal exposure*), gives insight into the distribution of dermal exposure over the body, and indicates by which routes dermal exposure takes place. Together with the ranking of tasks and jobs, this provides information for measurement strategies and helps to determine who, where, and what to measure (Van-Wendel-de-Joode et al., 2003).

When developing a semi-quantitative exposure assessment method, two important questions arise. First, are the results reliable? In other words, do different observers produce the same results for the same exposure situation? Second, are the estimates valid, do they represent the underlying real exposures? The aim of this study was to assess the reliability of DREAM by (i) studying inter-observer agreement, (ii) assessing the effect of individual observers on dermal exposure estimates for different tasks, and (iii) comparing inter-observer agreement with regard to the ranking of the exposure of body parts.

In order to study these questions, occupational hygienists and other occupational health professionals were asked to fill in DREAM while performing side-by-side observations for dermal exposure to a broad range of chemical agents for different tasks.

Methods

DREAM

In the following four paragraphs, we summarize the methodology of DREAM. For a detailed description of the method, we refer to Van-Wendel-de-Joode et al. (2003).

The basis for DREAM is the conceptual model for dermal exposure of Schneider and colleagues (1999), which describes how dermal exposure can occur, and thus by which routes substances can get onto the skin. The most important exposure routes are *emission*, *transfer*, and *deposition*, resulting in exposure on the outer clothing and skin. Schneider et al. (1999) defined *emission* as dermal exposure occurring directly from the source of exposure, *transfer* as exposure due to contact with contaminated surfaces, and *deposition* as exposure through skin contact with small particles present in the air compartment.

DREAM consists of two parts: a multiple-choice questionnaire (inventory part) on exposure determinants, and an evaluation algorithm. An occupational hygienist, or another occupational health professional, should fill in the questionnaire, after observing the worker(s). The questionnaire contains questions on dermal exposure determinants such as probability and intensity of dermal exposure routes (emission, transfer, and deposition); physical and chemical characteristics of the substance to which exposure occurs; percentage of working time task is being performed; and information on clothing layer (i.e., kind of material covering the skin, replacement frequency of clothing, percentage of task duration gloves are being worn). Each answer of the questionnaire coincides with a pre-assigned value that is subsequently put into the evaluation algorithm.

The algorithm comprises a systematical evaluation of determinants at the task level using default values that increase and decrease on a log scale, that is, 0.3, 1, 3, 10 as Cherrie and colleagues (1996) proposed for the evaluation of air-borne exposures. The directions of the default values of DREAM (increasing vs. decreasing exposure) are derived from the literature and expert judgment. The evaluation algorithms result in numerical estimates for exposure levels on both the outside clothing layer (*potential dermal exposure*) and, after taking into account the reductive effect of clothing, on the skin (*actual dermal exposure*). Exposure estimates are provided for nine individual body parts (hands, forearms, upper arms, head, torso front, back, lower abdomen/upper legs, lower legs, and feet) and the total body. The total body exposure on the outside clothing layer (total *potential* dermal exposure: $\text{Skin}_w{-}P_{\text{TASK}}$) and the skin (total *actual* dermal exposure: $\text{Skin}_w{-}A_{\text{TASK}}$) comprise the sum of individual body parts, after weighing body parts according to their surface area (Van-Wendel-de-Joode et al., 2003).

The numerical estimates are expressed in DREAM units, ranging from 0 to 40545 for the total body estimate. In addition to a numerical estimate, DREAM provides a categorical estimate for dermal exposure comprising seven ordinal categories defined by expert judgment (Van-Wendel-de-Joode et al., 2003).

Study Design

Four studies were performed in which a total of 29 observers (O) performed side-by-side observations for exposures to a broad range of liquid and solid chemical agents (e.g., insecticides, metal working fluids, solvents, chalk) for different tasks (Table 1). While observing each task, the observers filled in the multiple-choice questionnaire of the DREAM method. Two of the 29 observers performed observations in two studies.

The first study comprised seven tasks by three observers (O1, O2, and O3) who performed 56 side-by-side observations on grape farms in the Western Cape Province of South Africa. The seven tasks were mixing, pesticide spraying, horticulture summer tasks (trellising, thinning, and trimming of vines; connecting vines to wire; braiding of vines), water pipe maintenance, mist blower maintenance, manure spreading, and horticulture winter tasks (pruning of vines). The second study comprised two observers (O4 and O5) who made 24 side-by-side observations for 12 tasks at a truck manufacturing company, a rubber factory, and a hospital pharmacy in the Netherlands. The third study consisted of two to four observers (O2, O5, O6, and O7) who performed 21 side-by-side observations for 10 tasks in chemical companies in the Netherlands. The fourth study concerned 22 occupational hygienists who applied the method while observing two video-recorded tasks: a metal worker taking out connection rods of a CNC metalworking machine and a pharmacy assistant preparing the antineoplastic drugs (ANP) "cyclophosphamide". Each video-recorded task was played once. The observations of video recordings were included to see whether inter-observer agreement was comparable to observations at the workplace.

Training for all observers included (i) a short explanation of the conceptual model of Schneider and colleagues (1999) and the DREAM questionnaire, (ii) agreement with the observers on where tasks started and ended, and (iii) one to five practice runs to familiarize observers with the method's application.

Observers (O) had different backgrounds: O1 was a public health student, O3 was an occupational health nurse, while O2, O4, O5, O6 and O7 were occupational hygienists working at a research institute or a university. O8–O29 were practising occupational hygienists.

Statistical Analyses

Statistical analyses were performed with SAS System Software V8.2™. The distribution of DREAM dermal exposure estimates was tested and appeared to follow a lognormal distribution. All statistical analyses, except for the analyses on the DREAM categories, were therefore performed on log-transformed values. For each of the four studies, observers' intra-class correlation coefficients (ICCs) were calculated for *potential* and *actual* dermal exposure estimates, by dividing the variance explained by "task" by the total variance (Eq. (1)). The same coefficients were estimated for each of the three dermal exposure routes (*emission*, *transfer*, and *deposition*). Variance components were estimated by the mixed linear model procedure of SAS (PROC MIXED) using the method of restricted maximum likelihood (REML) and a compound symmetric variance matrix (CS).

$$\mathrm{ICC} = \frac{\sigma^2_{\mathrm{task}}}{\sigma^2_{\mathrm{total}}} \tag{1}$$

ICCs are a measure of inter-observer agreement for continuous data and may be interpreted as the Cohen's kappa statistics for categorical data (Fleiss 1975). Landis and Koch (1977) graded the levels of agreement as poor (less than 0.40), good (0.40–0.75), and excellent (≥ 0.75).

In order to assess the effect of individual observers on the *potential* and *actual* dermal exposure estimates for different tasks, mixed linear models (PROC MIXED using REML and CS) were applied for each of the four studies (Eq. (2)).

$$\begin{aligned} Y_{ij} &= \ln(X_{ij}) \\ &= \mu_j + \sum_{t=1}^{t} \beta_t \mathrm{task}_t + \sum_{c=1}^{c} \chi_c \mathrm{company}_c \\ &\quad + \sum_{tc=1}^{tc} \delta_{tc} \mathrm{task}_t \mathrm{company}_c + \omega_i + \varepsilon_{ij} \end{aligned} \tag{2}$$

where X_{ij} is the dermal exposure estimate (DREAM units) for the *j*th repetition for the *i*th observer; μ_j represents the true unknown mean (logged) dermal exposure estimate; β_t is the regression coefficient for task_t representing the effect of the *t*th task; χ_c is the regression coefficient for $\mathrm{company_c}$ (or department)[3] representing the effect of the *c*th company (or department); δ_{tc} is the regression coefficient for the *t*th task observed in the *c*th company (or department) representing the effect of the interaction of the *t*th task and the *c*th company (or department); ω_i is the random effect for the *i*th observer, which corresponds to the discrepancy between his/her intercept and overall intercept μ_j; and ε_{ij} is the random effect for residual variance.

In the models the variables "task" and "company/department" were included as *fixed effects*, while the variable "observer" was included as a *random effect* because we assumed individual observers to share a common mean exposure and to originate from the same distribution. To allow for differences in DREAM estimates for the same tasks in different companies/departments, an interaction term of "task" and "company/department" was included for the studies of grape farming (study 1), truck factory, rubber factory and pharmacy (study 2), and chemical industry (study 3), as in these studies some of the tasks were observed

3. In study 1 (grape farming) and study 3 (chemical industry), some tasks were observed repeatedly in different companies; in study 2 (truck factory, rubber factory, and pharmacy), some tasks were observed repeatedly in different departments.

Table 1. Geometric means (GMs) and range of total *potential* and *actual* dermal DREAM exposure estimates (in DREAM units) for each task.

Study	Observer code	Exposure to	Task	*N*[a]	Number of side-by-side observations	Number of companies or departments[b]	GM (GSD) of DREAM estimates	
							Potential dermal exposure	*Actual* dermal exposure
1. Grape farming	O1, O2, O3	Insecticides (organophosphates, carbamates)	Mixing	33	13	8	754 (6.2)	121 (5.3)
			Spraying	34	13	8	153 (2.3)	25 (2.8)
			Horticulture summer tasks	68	24	8	7 (1.9)	5 (1.7)
			Water pipe maintenance	4	2	1	1.5 (1.4)	0.4 (1.1)
			Mist blower maintenance	4	2	2	8 (2.4)	4 (2.8)
			Manure spreading	3	1	1	11 (1.6)	5 (1.4)
			Horticulture winter tasks	2	1	1	1 (1.2)	0.5 (1.0)
2. Truck, rubber factory, and pharmacy	O4, O5	Metal working fluids	Taking out MP of CNC[c] machines	12	6	2	13 (2.1)	4 (1.9)
			Taking out MP of other machines	10	5	2	39 (1.7)	8 (1.6)
		Solvents	Cleaning MP	2	1	1	96 (2.4)	16 (4.4)
			Spray painting	8	4	3	209 (2.7)	12 (1.9)
			Cleaning spray gun	2	1	1	268 (1.6)	34 (2.3)
			Unmasking truck parts	2	1	1	0 (1.0)	0 (1.0)
		Vulcanite (powder)	Filling (half automatic)	2	1	1	973 (1.5)	77 (2.5)
		Complex mixture (hot rubber)	Operating rubber machine (curing press)	2	1	1	30 (2.9)	10 (2.7)
		Complex mixture (warm rubber)	Operating rubber machine (open mill)	2	1	1	5 (4.4)	2 (3.1)
		Sulfur (filler)	Sealing bags	2	1	1	9 (1.5)	2 (1.5)
		Rubber chemicals	Manual dumping (pellets)	2	1	1	1 (1.8)	0 (1.0)
		Cyclophosphamide	Preparing ANP	2	1	1	2 (1.1)	0 (1.0)
3. Chemical industry	O2, O5, O6, O7	Chalk powder	Dumping (powder)	6	2	1	1251 (2.4)	156 (3.9)
		DEGBE[d]	Loading	11	4	4	14 (3.1)	4 (3.0)
			Filling (half automatic)	14	6	4	1 (2.7)	1 (2.1)
			Transferring by pumping	3	1	1	1 (1.5)	0 (1.1)
		Pigment (powder)	Filling (half automatic)	3	1	1	193 (3.3)	51 (3.0)

		Potassium hydroxide	Manual filling drums	4	2	1	21 (3.8)	6 (3.5)
		Petroleum	Wiping	6	2	1	41 (2.1)	11 (3.1)
		Solvents	Manual mixing paint	3	1	1	9 (1.1)	0.6 (1.2)
		Hydrochloride	Putting cans in wash machine	2	1	1	16 (1.0)	2 (1.1)
			Taking out cleaned cans	2	1	1	1 (1.5)	0 (1.1)
4. Video	O8 to O29	Metal working fluid	Taking out MP of CNC machine	22	1	1	59 (3.6)	41 (3.2)
		Cyclophosphamide	Preparing ANP	22	1	1	2 (1.9	1 (1.7)

[a] N = total number of observations.
[b] In study 1 (grape farming) and study 3 (chemical industry), some tasks were observed repeatedly in different companies; in study 2 (truck factory, rubber factory, and pharmacy), some tasks were observed repeatedly in different departments.
[c] CNC = computer numeric control.
[d] DEGBE = diethyl glycol butyl ether.

in different companies (studies 1 and 3) or departments (study 2) (Eq. 2). Geometric mean (GM) dermal exposure estimates for different tasks were consequently estimated by taking the exponent of the predicted betas for the tasks. In addition, the models estimated between-observer variance and whether observers' intercepts (ω_i) differed from the overall intercept (μ_j) (t-test $\alpha = 0.05$). By taking the exponent of the predicted observer's intercept ($\exp^{\omega i}$), a multiplier (M_O) was estimated for each observer. For the 22 observers of the video observations, median, 25th percentile, and 75th percentile observers' multipliers are presented. In order to get observer-specific estimates for tasks, the predicted GMs for tasks have to be multiplied by the observers' multipliers.

The model ascribed above was also applied with the variable *DREAM category* as Y-variable (Eq. (3)). Seven DREAM categories (0–6) exist for the dermal exposure estimates: 0 = no exposure (estimate = 0); 1 = very low exposure (estimates >0–10); 2 = low exposure (estimates 10–30); 3 = moderate exposure (estimates 30–100); 4 = high exposure (estimates 100–300); 5 = very high exposure (estimates 300–1000); and 6 = extremely high exposure (estimates >1000) (Van-Wendel-de-Joode et al., 2003). Consequently, the model predicted average DREAM categories for each task and individual observers' deviances (difference between an observer's intercept and overall intercept (μ_j) of the model).

$$Y_{ij} = \mu_j + \sum_{t=1}^{t} \beta_t \text{task}_t + \sum_{c=1}^{c} \chi_c \text{company}_c + \sum_{tc=1}^{tc} \delta_{tc} \text{task}_t \text{company}_c + \omega_i + \varepsilon_{ij} \quad (3)$$

where Y_{ij} is the DREAM dermal exposure category (0–6) for the jth repetition for the ith observer; μ_j represents the true unknown mean dermal exposure category; β_t is the regression coefficient for task$_t$ representing the effect of the tth task; χ_c is the regression coefficient for company$_c$ (or department) (see footnote 3) representing the effect of the cth company (or department); δ_{tc} is the regression coefficient for the tth task observed in the cth company (or department) representing the effect of the interaction of the tth task and the cth company (or department); ω_i is the random effect for the ith observer, which corresponds to the observer's deviance: the difference between an observer's intercept and overall intercept (μ_j) of the model; and ε_{ij} is the random effect for residual variance.

Inter-observer agreement for ranking of dermal exposure of body parts was studied by estimating Spearman rank correlation coefficients (PROC CORR Spearman) for each pair of observers. For the study of the video observations, the median value of the correlation coefficients of all observers was presented, as it comprised two tasks observed by 22 observers.

Table 2. Intra class correlation coefficients (ICCs) for *potential* and *actual* dermal DREAM exposure estimates, and dermal exposure route estimates: *emission*, *transfer*, and *deposition* for each of the four studies.

Study	Total observations (N^a)	Tasks (n^b)	Observers (K^c)	ICCs of dermal DREAM exposure estimates				
				Potential	Actual	Emission	Transfer	Deposition
1. Grape farming	148	7	3	0.79	0.68	0.80	0.61	0.74
2. Truck, rubber factory, pharmacy	48	12	2	0.87	0.78	0.72	0.78	0.79
3. Chemical industry	54	10	4	0.81	0.72	0.74	0.71	0.72
4. Video (truck factory, pharmacy)	44	2	22	0.81	0.83	0.87	0.48	0.69

[a]N = total number of observations.
[b]n = number of observed tasks.
[c]K = number of observers.

Results

Table 1 presents Geometric means (GMs) and geometric standard deviations (GSDs) of total *potential* and *actual* dermal exposure estimates for each of the 31 observed tasks. Observed tasks covered a wide range of DREAM estimates, falling in the DREAM categories of no (estimate = 0) to extremely high (estimate >1000) dermal exposure (Van-Wendel-de-Joode et al., 2003). Overall, GMs for tasks ranged from 0 (unmasking truck parts) to 1251 (dumping powder) for *potential* dermal exposure estimates, while GMs for *actual* estimates ranged from only 0 to 156.

Table 2 presents ICCs between observers for both total *potential* and *actual* dermal exposure estimates, as well as for the dermal exposure route estimates: *emission*, *transfer*, and *deposition*. ICCs ranged from 0.48 to 0.87. ICCs of total *potential* DREAM estimates were highest for the second study and practically identical for the other three studies. ICCs of total *actual* exposure estimates were somewhat lower than *potential* estimates, except for the video observations.

Agreement for the different exposure routes varied between the four studies. The first (grape farming) and fourth studies (video) showed highest agreement for *emission* (0.80 and 0.87, respectively), while agreement for *transfer* was lowest (0.61 and 0.48, respectively). Probably, observers had difficulties with assessing contamination levels of surfaces by means of watching a video, explaining the relatively low inter-observer agreement (0.48) for transfer for the video-taped situations. In grape farming, observers may have encountered difficulties assessing surface contamination for horticulture tasks as information on the date of last pesticide application and on pesticide degradation was lacking. For the second study (truck, rubber company, pharmacy), *transfer* and *deposition* showed somewhat higher agreement (0.78 and 0.79, respectively) as compared to *emission* (0.72). The third study (chemical industry) showed practically identical agreement for the three exposure routes (0.74, 0.71, and 0.72 for *emission*, *transfer*, and *deposition*, respectively).

Table 3 shows model predictions of GMs (with 95% confidence intervals, CI) of *potential* and *actual* dermal exposure estimates for each task, and predicted multipliers for each observer (M_O). In addition, estimates of between observer (S_o^2) and residual variance (S_e^2) are presented. In all cases the observers' multipliers did not differ statistically significantly from 1 (t-test $\alpha = 0.05$). Differences in observers' dermal exposure estimates were largest for the first study of grape farming, where observers' multipliers were 0.5, 1 and 1.9 for *potential*, and 0.7, 1.0 and 1.6 for *actual* dermal exposure estimates. Observers' multipliers were smallest for the third study (chemical industry), ranging from 0.9 to 1.2 and 0.9 to 1.1 for *potential* and *actual* dermal exposure estimates, respectively. Except for the video observations (study 4), differences in observers' estimates were somewhat lower for *actual* dermal exposure estimates as compared to *potential* dermal exposure estimates.

All four studies showed statistically significant differences ($P < 0.0001$) in predicted GMs of *potential* and *actual* dermal exposure estimates for tasks (Table 3). Predicted GMs ranged from 0 to 1226 for potential dermal exposure estimates, and 0 to 154 for actual dermal exposure estimates. In addition, differences in dermal exposure between companies and departments were found, and for one study (grape farming) DREAM was able to detect differences in dermal exposures for the same task performed in different farms (data not shown).

Table 4 presents predictions of averages of *potential* and *actual* DREAM dermal exposure categories for each task, and predicted mean observers' deviation. None of the predicted observer deviations differed significantly from 0 (t-test $\alpha = 0.05$). Predicted mean observer deviations ranged from -0.42 to 0.40 dermal exposure category, and were largest for study 1 (grape farming). Again, observers' mean deviations were smaller for *actual* dermal exposure estimates, except for the study of the video observations (study 4). Tasks showed significant differences in average dermal exposure categories ($P < 0.0001$) for both *potential* and *actual* exposures; the predicted averages ranged from 0.0 to 5.5.

Inter-observer agreement for the ranking of dermal exposure of nine body parts was estimated for each pair of observers for each exposure situation with Spearman rank

Table 3. Model predictions of GM with 95% CI of *potential* and *actual* dermal exposure estimates for each task (fixed effect)[a], predicted multipliers for each observer (M_O) (random effect) in DREAM units, and predicted between-observer (S_o^2) and residual variance (S_e^2) (random effects).

Study	Potential dermal exposure estimates			Actual dermal exposure estimates		
Task*	GM (95% CI)	M_O (95% CI)	Variance components	GM (95% CI)	M_O (95% CI)	Variance components
1. Grape farming						
Mixing	560 (260 – 1206)	O1: 1.9 (0.9 – 3.9)	$S_o^2=0.39$	117 (65 – 208)	O1: 1.6 (0.9 – 2.7)	$S_o^2=0.21$
Spraying	149 (68 – 322)	O2: 0.5 (0.3 – 1.1)	$S_e^2=0.56$	28 (15 – 51)	O2: 0.7 (0.4 – 1.1)	$S_e^2=0.43$
Horticulture summer tasks	7 (3 – 16)	O3: 1.0 (0.5 – 2.1)		5 (2– 9)	O3: 1.0 (0.6 – 1.7)	
Water pipe maintenance	1 (0 – 6)			0 (0 – 2)		
Mist blower maintenance	9 (3 – 28)			4 (1 – 12)		
Manure spreading	11 (3 – 35)			5 (1 – 14)		
Horticulture winter tasks	2 (0 – 8)			1 (0 – 4)		
2. Truck, rubber, pharmacy						
Taking out MP of CNC machines[b]	17 (7 – 41)	O4: 1.4 (0.7 – 2.8)	$S_o^2=0.23$	5 (2 –11)	O4: 1.3 (0.7 – 2.5)	$S_o^2=0.18$
Taking out MP of other machines	37 (16 – 87)	O5: 0.7 (0.4 – 1.5)	$S_e^2=0.34$	7 (3 – 16)	O5: 0.7 (0.4 – 1.4)	$S_e^2=0.29$
Cleaning MP	96 (32 – 286)			16 (5 – 46)		
Spray painting	168 (73 – 385)			12 (5 – 25)		
Cleaning spray gun	268 (90 – 797)			34 (12– 95)		
Unmasking	0 (0 – 2)			0 (0 – 2)		
Filling (half automatic)	973 (327 – 2890)			77 (28 – 210)		
Curing press	30 (10 – 92)			10 (3 – 30)		
Open mill	5 (1 – 16)			2 (0 – 8)		
Sealing bags	9 (3 – 30)			2 (0 – 7)		
Manual dumping	1 (0 – 3)			0 (0 – 2)		
Preparing ANP[c]	2 (0 – 7)			0 (0 – 2)		
3. Chemical industry						
Dumping	1226 (613 – 2452)	O2: 1.0 (0.7 – 1.4)	$S_o^2=0.04$	154 (70 – 335)	O2: 1.0 (0.7 – 1.4)	$S_o^2=0.02$
Loading	13 (7 – 24)	O5: 1.2 (0.8 – 1.7)	$S_e^2=0.63$	3 (1 – 7)	O5: 1.1 (0.8 – 1.5)	$S_e^2=$ 0.82
Filling	2 (1 – 4)	O6: 1.0 (0.7 – 1.5)		1 (0 – 3)	O6: 1.0 (0.7 – 1.4)	
Filling (powder)	190 (72 – 497)	O7: 0.9 (0.6 – 1.2)		51 (17 – 153)	O7: 0.9 (0.6 – 1.2)	
Manual filling drums	21 (8 – 50)			6 (2 – 18)		
Wiping	41 (20 – 83)			11.9 (5.5 – 26.0)		
Manual mixing paint	9 (3 – 25)			1 (0 – 4)		
Transferring	1 (0 – 4)			0 (0 – 3)		
Putting cans	16 (4 – 54)			2 (0 – 12)		
Taking out cans	1 (0 – 6)			0 (0 – 4)		
4. Video						
Taking out MP of CNC machines	59 (37 – 92)	O_{50}: 0.8 (0.4 – 1.9)[d]	$S_o^2=0.22$	41 (27 – 62)	O_{50}: 0.9 (0.4 – 2.0)	$S_o^2=0.23$
Preparing ANP	2 (1 – 4)	O_{25}: 0.6 (0.4 – 2.1) O_{75}: 1.1 (0.5 – 2.5)	$S_e^2=0.80$	1 (1 – 3)	O_{25}: 0.8 (0.4 – 1.6) O_{75}: 1.2 (0.6 – 2.5)	$S_e^2=0.61$

[a]In addition to "task", the fixed effects "company/department" and an interaction term of "task" with "company/department" were included in the model. Predicted dermal exposure levels for each task were adjusted for company/department and interaction between task and company/department.
[b]MP = metal parts, CNC = computer numeric control.
[c]ANP = antineoplastic drugs.
[d]O_{50} = observer with median factor; O_{25} = observer with 25th percentile factor; O_{75} = observer with 75th percentile factor.
*The effect of "task" was statistically significant ($P<0.0001$) for all four studies.

Table 4. Model predictions of average DREAM exposure category (0–6) with 95% CI of exposure on clothing and skin for each task (fixed effect)[a], predicted mean deviation for each observer (D_O) (random effect), and predicted between-observer (S_o^2) and residual variance (S_e^2) (random effects).

Study	Potential dermal exposure category			Actual dermal exposure category		
Task*	Mean (95% CI)	D_O (95%CI)	Variance components	Mean (95% CI)	D_O (95%CI)	Variance components
1. Grape farming						
Mixing	4.7 (4.1 – 5.2)[b]	O1: 0.40 (−0.10 – 0.89)	$S_o^2=0.18$	3.6 (3.2 – 4.0)	O1: 0.33 (−0.05 – 0.72)	$S_o^2=0.10$
Spraying	3.9 (3.4 – 4.4)	O2: −0.42 (−0.92 – 0.07)	$S_e^2=0.30$	2.4 (2.0 – 2.8)	O2: −0.24 (−0.62 – 0.14)	$S_e^2=0.37$
Horticulture summer tasks	1.2 (0.7 – 1.7)	O3: 0.02 (−0.47 – 0.52)		1.0 (0.6 – 1.4)	O3: −0.09 (−0.48– 0.30)	
Water pipe maintenance	1.0 (0.3 – 1.7)			1.0 (0.3 – 1.7)		
Mist blower maintenance	1.4 (0.6 – 2.1)			1.3 (0.6 – 2.0)		
Manure spreading	1.7 (0.9 – 2.5)			1.0 (0.2 – 1.8)		
Horticulture winter tasks	1.2 (0.3 – 2.1)			1.2 (0.2 – 2.1)		
2. Truck, rubber, pharmacy						
Taking out MP of CNC mach.[c]	2.2 (1.5 – 28)	O4: 0.25 (0.30– 0.80)	$S_o^2=0.13$	1.1 (0.5 –1.6)	O4: 0.16 (−0.22 – 0.54)	$S_o^2=0.06$
Taking out MP of other mach.	2.6 (1.9 – 3.2)	O5: −0.25 (−0.80 – 0.30)	$S_e^2=0.30$	1.3 (0.8 – 1.8)	O5: −0.16 (−0.54 – 0.22)	$S_e^2=0.25$
Cleaning MP	3.5 (2.6 – 4.4)			2.0 (1.2 – 2.8)		
Spray painting	3.9 (3.2 – 4.6)			1.6 (1.1 – 2.1)		
Cleaning spray gun	4.5 (3.6 – 5.4)			2.5 (1.7– 3.3)		
Unmasking	0.0 (−0.9 – 0.9)			0.0 (−0.8 – 0.8)		
Filling (half automatic)	5.5 (4.6 – 6.4)			3.5 (2.7 – 4.3)		
Curing press	2.5 (1.6 – 3.4)			1.5 (0.7 – 2.3)		
Open mill	1.5 (0.6 – 2.4)			1.0 (0.2 – 1.8)		
Sealing bags	1.5 (0.6 – 2.4)			0.5 (−0.3 – 1.3)		
Manual dumping	1.5 (0.5 – 4.5)			1.0 (0.2 – 1.8)		
Preparing ANP	2.8 (0.9 – 8.2)			1.0 (0.2 – 1.8)		
3. Chemical industry						
Dumping	5.3 (4.8 – 5.9)	O2: −0.02 (−0.33 – 0.29)	$S_o^2=0.03$	3.8 (3.3 – 4.4)	O2: 0.0 (.. - ..)[e]	$S_o^2=0.00$
Loading	1.9 (1.5 – 2.3)	O5: 0.11 (−0.16 – 0.38)	$S_e^2=0.40$	1.3 (0.8 – 1.7)	O5: 0.0 (.. - ..)	$S_e^2=0.42$
Filling	1.2 (0.8 – 1.6)	O6: 0.05 (−0.24 – 0.34)		1.1 (0.7 – 1.4)	O6: 0.0 (.. - ..)	
Filling (powder)	4.0 (3.2 – 4.8)	O7: −0.14 (−0.41 – 0.13)		3.0 (2.2 – 3.8)	O7: 0.0 (.. - ..)	
Manual filling drums	2.3 (1.6 – 2.9)			1.5 (0.8 – 2.2)		
Wiping	2.8 (2.3 – 3.4)			1.8 (1.3 – 2.4)		
Manual mixing paint	1.0 (0.3 – 1.8)			1.0 (0.2 – 1.8)		
Transferring	1.0 (0.3 – 1.8)			1.0 (0.2 – 1.8)		
Putting cans	2.0 (1.1 – 2.9)			1.0 (0.1 – 1.9)		
Taking out cans	1.0 (0.1 – 1.9)			1.0 (0.1 – 1.9)		
4. Video						
Taking out MP of CNC mach.[c]	3.1 (2.7 – 3.4)	O_{50}: 0.00 (−0.14 – 0.14)[d]	$S_o^2=0.00$	2.8 (2.5 – 3.1)	O_{50}: 0.02 (−0.33 – 0.36)	$S_o^2=0.03$
Preparing ANP	1.0 (0.6 – 1.3)	O_{25}: −0.01 (−0.14 – 0.13)	$S_e^2=0.63$	0.9 (0.6 – 1.2)	O_{25}: −0.04 (−0.38 – 0.30)	$S_e^2=0.47$
		O_{75}: 0.01 (−0.13 – 0.14)			O_{75}: 0.02 (−0.33 – 0.36)	

[a]In addition to "task", the fixed effects "company/department" and an interaction term of "task" with "company/department" were included in the model. Predicted dermal exposure categories for each task were adjusted for company/department and interaction between task and company/department.
[b]ANP = antineoplastic drugs.
[c]MP = metal parts, CNC = computer numeric control, mach = machines.
[d]O_{50} = observer with median factor; O_{25} = observer with 25th percentile factor; O_{75} = observer with 75th percentile factor.
[e]Confidence intervals were not estimable.
*The effect of "task" was statistically significant ($P<0.0001$) for all four studies.

Table 5. Median value of Spearman correlation coefficients for ranking of *potential* and *actual* dermal exposure estimates of nine body parts for pairs of observers for each study.

Study	Median	
	Potential dermal exposure estimates	*Actual* dermal exposure estimates
1. Grape farming		
O1 – O2 ($n=52$)	0.49	0.65
O1 – O3 ($n=36$)	0.29	0.49
O2 – O3 ($n=40$)	0.49	0.58
2. Truck, rubber factory, pharmacy		
O4 – O5 ($n=22$)	0.81	0.70
3. Chemical industry		
O2 – O5 ($n=5$)	0.68	0.87
O2 – O7 ($n=3$)	0.99	0.72
O5 – O6 ($n=8$)	0.96	0.93
O5 – O7 ($n=16$)	0.74	0.81
O6 – O7 ($n=10$)	0.75	0.68
4. Video		
All observers ($n=392$)[a]	0.83	0.92

[a]Median value of the correlation coefficients of all observers is presented.

correlation coefficients (PROC CORR spearman) for *potential* and *actual* dermal exposure estimates (Table 5). Median correlations of pairs of observers ranged from 0.29 to 0.99, and were lowest for the first study (grape farming). Except for the second study (truck and rubber factory, pharmacy), the ranking of body parts was more similar for *actual* dermal exposure estimates as compared to *potential* dermal exposure estimates.

Discussion

We studied inter-observer agreement of DREAM for a broad range of tasks with mainly exposures to liquids, but also with exposure to solids, vapors, and fumes.

Inter-observer agreement was studied by estimating ICCs. In our studies ICCs between observers ranged from 0.68 to 0.83 for *actual*, and 0.79 to 0.87 for *potential* dermal exposure estimates, indicating good and excellent inter-observer agreement (Landis and Koch, 1977). Except for the study of video observations, *actual* dermal exposure estimates showed somewhat lower inter-observer agreement than *potential* dermal exposure estimates. *Actual* dermal exposure estimates include a *clothing protection factor* containing an additional source of disagreement. With regard to the video observations, part of the information on clothing was provided beforehand, which explains that ICCs for *potential* and *actual* dermal exposure estimates were nearly the same.

ICCs for the different exposure routes — *emission*, *transfer*, and *deposition* — indicated a good to excellent inter-observer agreement as well, as they ranged from 0.48 to 0.87. In general, *emission* showed highest inter-observer agreement followed by *deposition* and *transfer*. Agreement between observers for the different exposure routes varied slightly between studies, probably due to differences in tasks. The occurrence of exposure routes varies between tasks, and for some tasks it may be easier to assess frequency and intensity of occurrence of an exposure route than for other tasks. However, the differences in inter-observer agreement could also be due to observers, as they varied between studies as well; only two of the 29 observers performed observations in two studies.

Also when compared to other studies, the DREAM method showed good inter-observer agreement. Kromhout et al. (1987) reported kappa values between 0.23 and 0.50 for two occupational hygienists, who ranked tasks with inhalation exposures to total solvents, and different types of dust (organic, fibers, welding fumes) on a semi-quantitative 4-point scale by means of workplace observations. Hertzman et al. (1988) found an ICC of 0.55 for 10 workers, applying an exposure index that was based on assessed frequency and duration of exposure to chlorophenates for 59 jobs. Siematycki et al. (1997) reported a weighted kappa of 0.70 for two experts assessing inhalation exposures for 294 workplace chemicals on the basis of job histories. In a study by Benke et al. (1997) five experts rated exposure to 21 chemicals for 298 job descriptions. The three, of a total of 21 chemicals (cutting fluids, welding fumes, lubricating oils), with greatest inter-rater agreement had weighted kappas between 0.42 and 0.64 (Benke et al., 1997). Correlation coefficients for log-transformed values of quantitative estimates made by five occupational hygienists, assessing inhalation exposures to asbestos, man-made vitreous fibers, dust, and polycyclic aromatic hydrocarbons using information supplied by the authors, ranged from 0.73 to 0.83 (Semple et al., 2001). de Cock et al. (1996) asked five occupational hygienists to rank dermal and inhalation exposure to pesticides for 14 tasks performed in fruit growing. The hygienists watched a video and received additional information. The ICCs for the five occupational hygienists were 0.72 for dermal, and 0.62 for inhalation exposure to pesticides (de Cock et al., 1996).

The effects of observers on the task estimates were relatively small, as the maximum predicted mean individual observer multiplier was only a factor 2 while predicted GMs of dermal exposure estimates for tasks ranged from 0 to 1226, and none of the predicted observers' multipliers significantly differed from 1. The largest individual observer multipliers were found in the first study (grape farming), possibly because these observers had different background and experiences in exposure assessment; observer 1 was a public health student, observer 2 an occupational hygienist, whereas observer 3 was an occupational nurse. The observers of the other studies were all occupational hygienists. However, observing different tasks in different studies could also have caused differences; it could have been more difficult for observers to fill in DREAM for certain grape farming tasks than for the tasks studied in the other sub-studies covering miscellaneous industries and situations. The effect of observers on the estimated DREAM dermal exposure categories for each task was also small. The maximum predicted mean observers' deviation was less than half a category, whereas for tasks the predicted averages of the categories ranged from 0 to 5.5.

Inter-observer agreement for ranking of dermal exposure of nine body parts by applying DREAM was acceptable for *potential* exposure estimates, median values of Spearman correlation coefficients ranged from 0.29 to 0.83, and good for *actual exposure estimates* with median values from 0.49 to 0.93. Again, inter-observer agreement was lowest for grape farming.

As expected, DREAM was able to detect differences in dermal exposure because for all the four studies the effect of "task" on *potential* as well as *actual* dermal exposure was apparent. Underlying factors like actual task content, physical–chemical properties of the substance to which exposure occurs, and use of protecting clothing are reflected in the estimates. Dermal exposure estimates for tasks were considerably lower for *actual* dermal exposure than for *potential* dermal exposure, showing the effect of the *clothing protective factor*.

Conclusion

The systematical approach of DREAM, based on the conceptual model of Schneider and colleagues (1999), results in reproducible dermal exposure estimates and reduces the effect of observers' subjectivity. However, it is still unclear whether DREAM estimates reflect absolute dermal exposure levels or rather relative dermal exposure estimates within workplaces.

DREAM is a method that can be applied reliably to estimate *potential* or *actual* dermal exposure to a broad range of chemical agents for distinct tasks in all kinds of industries.

DREAM appears to offer a useful advance for estimations of dermal exposure both for epidemiological research and for occupational hygiene practice.

Acknowledgements

We thank Simone Hilhorst, Martijn Kerkman, Mark Lurvink, Motlagomang Maruping, and occupational hygienists of the eastern part of the Netherlands for observing exposure situations and applying DREAM. We are grateful to Jim Deddens and Brian Miller for their statistical advice. This study was financed by the Dutch Ministry of Social Affairs and Employment and partly facilitated by the South Africa Netherlands Research Program on Alternatives in Development 00/17(SANPAD). The cooperation of employers and employees who participated in the study is gratefully acknowledged.

References

Benke G., Sim M., Forbes A., and Salzberg M. Retrospective assessment of occupational exposure to chemicals in community-based studies: validity and repeatability of occupational hygiene panel ratings. *Int J Epidemiol* 1997: 26(3): 635–642.

Brouwer D.H., Semple S., Marquart J., and Cherrie J.W. A dermal model for spray painters. Part I: subjective exposure modelling of spray paint deposition. *Ann Occup Hyg* 2001: 45(1): 15–23.

Cherrie J.W., and Schneider T. Validation of a new method for structured subjective assessment of past concentrations. *Ann Occup Hyg* 1999: 43(4): 235–245.

Cherrie J.W., Schneider T., Spankie S., and Quinn M. A new method for structured, subjective assessments of past concentrations. *Occup Hyg* 1996: 3: 75–83.

Cherrie J.W., Tickner J., Friar J., Rae R., Warren N., and Pryde D.E. Evaluation and further development of the EASE model 2.0. Research report 136, prepared by the Institute of Occupational Medicine for the Health and Safety Executive, ISBN 0 7176 2714 4, 2003.

de Cock J., Kromhout H., Heederik D., and Burema J. Experts' subjective assessment of pesticide exposure in fruit growing. *Scand J Work Environ Health* 1996: 22(6): 425–432.

Fleiss J.L. Measuring agreement between two judges on the presence or absence of a trait. *Biometrics* 1975: 31(3): 651–659.

Hawkins N.C., and Evans J.S. Subjective estimation of toluene exposures: a calibration study of occupational hygienists. *Appl Ind Hyg J* 1989: 4: 61–68.

Hertzman C., Teschke K., Dimich-Ward H., and Ostry A. Validity and reliability of a method for retrospective evaluation of chlorophenate exposure in the lumber industry. *Am J Ind Med* 1988: 14(6): 703–713.

Kromhout H. Commentary. *Occup Environ Med* 2002: 59: 594.

Kromhout H., Oostendorp Y., Heederik D., and Boleij J.S. Agreement between qualitative exposure estimates and quantitative exposure measurements. *Am J Ind Med* 1987: 12(5): 551–562.

Landis J.R., and Koch G.G. The measurement of observer agreement for categorical data. *Biometrics* 1977: 33(1): 159–174.

Post W., Kromhout H., Heederik D., Noy D., and Smit Duijzentkunst R. Semiquantitative estimates of exposure to methylene chloride and styrene: the influence of quantitative exposure data. *Appl Occup Environ Hyg* 1991: 6(3): 197–204.

Schneider T., Vermeulen R., Brouwer D.H., Cherrie J.W., Kromhout H., and Fogh C.L. Conceptual model for assessment of dermal exposure. *Occup Environ Med* 1999: 56: 765–773.

Semple S.E., Proud L.A., Tannahill S.N., Tindall M.E., and Cherrie J.W. A training exercise in subjectively estimating inhalation exposures. *Scand J Work Environ Health* 2001: 27(6): 395–401.

Siemiatycki J., Fritschi L., Nadon L., and Gerin M. Reliability of an expert rating procedure for retrospective assessment of occupational exposures in community-based case-control studies. *Am J Ind Med* 1997: 31(3): 280–286.

Teschke K., Hertzman C., Dimich-Ward H., Ostry A., Blair J., and Hershler R. A comparison of exposure estimates by worker raters and occupational hygienists. *Scand J Work Environ Health* 1989: 15(6): 424–429.

Teschke K., Olshan A.F., Daniels J.L., De Roos A.J., Parks C.G., and Schulz, et al. Occupational exposure assessment in case–control studies: opportunities for improvement. *Occup Environ Med* 2002: 59(9): 575–593.

Van-Wendel-de-Joode B., Brouwer D.H., Vermeulen R., Van Hemmen J.J., Heederik D., and Kromhout H. DREAM: a method for semi-quantitative dermal exposure assessment. *Ann Occup Hyg* 2003: 47(1): 71–87.

Vermeulen R., Stewart P., and Kromhout H. Dermal exposure assessment in occupational epidemiologic research. *Scand J Work Environ Health* 2002: 28(6): 371–385.

Walker K.D., Catalano P., Hammit J.K., and Evans J.S. Use of expert judgement in exposure assessment: Part 2. Calibration of expert judgements about personal exposures to benzene. *J Expo Anal Environ Epidemiol* 2003: 13: 1–16.